Humanities Data

Monday 4th August		
Time	Session	Speakers
09:00-10:30	**Opening keynote** See keynotes page	
10:30-11:00	**Refreshment break**	
11:00-11:30	**Introduction** The workshop convenors will give attendees an overview of the week's programme.	Neil Jefferies Meriel Patrick Rowan Wilson
11.30 - 12.30	**A critical review of humanities data approaches** Choosing the correct approach for your data can have a significant impact on the success, or otherwise, of your research, and this talk will encourage you to critically evaluate all standards and practices much as you would evaluate your scholarly sources.	Neil Jefferies
12:30-14:00	**Lunch**	
14:00-15:30	**Corpus linguistics** Before the terms 'big data', 'data science', 'evidence-based research', 'digital humanities', 'distant reading' or 'text mining' were current, there was 'corpus linguistics', in which a digital, data-driven approach was developed to text analysis and the study of language	Martin Wynne

	usage. Whether it is considered as a set of methods or an academic sub-discipline, corpus linguistics continues to thrive and grow as an approach to research in linguistics and related fields. Nowadays, there are a wide range of tools - applications, interfaces, environments and platforms of various types which are available to do corpus linguistics. This presentation will offer a tour of these tools, what you can and can't do with them, and offer some suggestions and tips on how to choose among the sometimes bewildering array of options.	
15.30-16.00	**Refreshment break**	
16:00-17:00	**Corpus linguistics** Continuation of earlier session.	Martin Wynne

Tuesday 5th August		
Time	Session	Speakers
09:00-10:30	**Data visualisation for the humanities** How can visualisation be used to explore and communicate humanities data? Using examples from research and collections, we will discuss visualisation design principles and give an overview of data visualisation technology.	Alexander Shiarella
10:30-11:00	**Refreshment break**	

11:00-12:30	**The Quill Project** The Quill Project is an interdisciplinary digital-humanities project that examines the negotiation of legal texts. This session will discuss problems of representing these kinds of processes, designing interfaces for data-entry, collaboration with undergraduate students, and visualisation design.	Nicholas Cole
12:30-14:00	**Lunch**	
14:00-15:30	**Digital preservation** Digital preservation involves protecting and maintaining digital content over time to ensure its long-term accessibility. The aim is to keep digital information available for future generations, even as technology evolves. This session will cover key digital preservation concepts and methods, including format conversion and emulation.	Edith Halvarsson
15:30-16:00	**Refreshment break**	

16:00-17:00	**Digital preservation: Optional hands-on session** Using the preservation of videogames as a window into some of the key topics around digital preservation and the history of technology, this hands-on session will invite participants to engage with a selection of items from the Digital Preservation Collection's archive of games, gaming peripherals, and supplementary material. Test your reflexes in *Chuckie Egg* (A&F Software, 1983), marvel at the pioneering vector graphics and immersive simulation of *Elite* (Acornsoft, 1984), and puzzle at the post-Monty Python surrealism of *Jet Set Willy* (Matthew Smith, 1984), all playable on our Acorn BBC Micro computer (1981). Browse a selection of magazines, player guides, and in-box ephemera from the 1980s and early 90s, and learn about the intersection of digital and non-digital preservation in capturing the history of videogames and the unique challenges in engaging with a medium which combines product, art form, experience, technology, and time capsule, all in one complex package.	

Wednesday 6th August		
Time	Session	Speakers
09:00-10:30	**Framing digital objects with context and provenance** Cultural and historical objects derive a lot of their meaning and interpretation from the contexts in which they are created and subsequently experienced. When digital surrogates or born-digital artefacts are created, it is important that this contextual information is also represented in the digital domain. This talk will explore the nature of context and provenance (which can be seen as a historical series of contexts), and consider how they might be modelled digitally.	Neil Jefferies
10:30-11:00	**Refreshment break**	

Time	Session	Speakers
11:00-12:30	**Framing digital objects with context and provenance (continued)** Continuation of the previous session.	Neil Jefferies
12:30-14:00	**Lunch**	
14:00-15:30	**IIIF - the International Image Interoperability Framework** With the increasing volume of digitised materials being made available via the International Image Interoperability Framework (IIIF), it is becoming an integral part of digital humanities workflows. This session will introduce you to the basics of IIIF and provide some examples of how it can be combined with other digital tools in such workflows.	Neil Jefferies
15:30-16:00	**Refreshment break**	
16:00-17:00	**Optional hands-on session** An optional opportunity for participants to explore the framework covered in the previous session.	Neil Jefferies

Thursday 7th August		
Time	Session	Speakers

09:00-10:30	**OpenRefine for tabular data** OpenRefine is a powerful tool for working with messy data. It can help you explore large datasets, clean them, and transform them from one format into another. Through hands-on exercises, this session will introduce you to the core features of OpenRefine.	Edith Halvarsson
10:30-11:00	**Refreshment break**	
11:00-12:30	**Introduction to relational databases** This session looks at what a relational database is, and when and why it might be helpful to use one. It introduces some basic database concepts, and works through the process of planning a relational database - with some hands-on exercises to help start putting the theory into practice. We also look at some challenges posed by the sort of data often used in humanities projects, and how these might be addressed.	Meriel Patrick Pamela Stanworth
12:30-14:00	**Lunch**	
14:00-15:30	**Using TEI to encode and extract data** This session will serve as an introduction to the principles and practice of TEI-XML, the most widespread scholarly method of encoding texts. We will look both at how to encode texts within TEI, and the kinds of data we can extract from it, including getting hands-on with some encoding.	Sebastian Dows-Miller
15:30-16:00	**Refreshment break**	
16:00-17:00	**Optional hands-on session**	Sebastian Dows-Miller

	An optional opportunity for participants to explore the techniques covered in the previous session.	

Friday 8th August

Time	Session	Speakers
09:00-10:30	**Computer vision and machine learning** Abstract: This session will explore computer vision - or the extraction of data from images. Although digital humanists have a rich set of tools for searching and analysing text, images have until recently been less amenable to processing. The session will showcase tools developed by Oxford's Visual Geometry Group in collaboration with researchers in the humanities working in fields including book and art history, film, music and photography. Participants will gain an understanding of the state-of-the-art in such computer vision operations as image comparison, classification and search; of how humanists are currently using these methods; and how they can apply them to their own materials. We will use web demos and (optionally) user-installable software. No coding is required.	Giles Bergel
10:30-11:00	**Refreshment break**	

11:00-12:30	**Humanities data miscellany** Short presentations on a range of topics. Attendees will be given the chance to talk about their work (or just a topic they think is really cool). This is your opportunity to share what you're doing, and to tap into the knowledge and experience of your fellow delegates.	Strand participants and/or convenors
12:30-14:00	**Lunch**	
14:00-15:00	**Wrap-up session** An opportunity to draw together the threads of the week's presentations, and to reflect on next steps. There will also be an opportunity to ask any final questions, and to draw on the insight and experience of your fellow participants.	The convenors
15:00-15:30	**Refreshment break**	
15.30-17.00	**Closing keynote** See keynotes page	